

P. LESAIN

P. A. RAVIART

**On a Finite Element Method for Solving the Neutron Transport Equation**

*Publications des séminaires de mathématiques et informatique de Rennes*, 1974, fascicule S4

« Journées éléments finis », , p. 1-40

[http://www.numdam.org/item?id=PSMIR\\_1974\\_\\_S4\\_A8\\_0](http://www.numdam.org/item?id=PSMIR_1974__S4_A8_0)

© Département de mathématiques et informatique, université de Rennes, 1974, tous droits réservés.

L'accès aux archives de la série « Publications mathématiques et informatiques de Rennes » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

ON A FINITE ELEMENT METHOD FOR SOLVING  
THE NEUTRON TRANSPORT EQUATION

*by* P. LESAIN<sup>\*</sup> and P.A. RAVIART<sup>\*\*</sup>

<sup>\*</sup> Service de Mathématiques Appliquées, Centre d'Etudes de Limeil, B.P. 27,  
94190 VILLENEUVE-SAINT-GEORGES FRANCE.

<sup>\*\*</sup> Analyse Numérique, Tour 55-65, Université PARIS VI , 4 Place Jussieu,  
75230 PARIS CEDEX 05 , FRANCE.

Paper presented at the Symposium on Mathematical Aspects of Finite Elements  
in Partial Differential Equations, Mathematics Research Center, University  
of Wisconsin, Madison, April 1.3, 1974.

1. - INTRODUCTION.

Let  $\Omega$  be a convex open set in the  $(x,y)$ -plane with boundary  $\Gamma$ . Denote by  $n_x$  and  $n_y$  the components of the outward unit vector normal to  $\Gamma$ .

Let  $Q$  be the unit disk in the  $(\mu,\nu)$ -plane. We consider the following problem : Find a function  $u = u(x,y,\mu,\nu)$  such that

$$(1.1) \quad \mu \frac{\partial u}{\partial x} + \nu \frac{\partial u}{\partial y} + \sigma u = f \quad \text{in} \quad \Omega \times Q ,$$

$$(1.2) \quad u(x,y,\mu,\nu) = 0 \quad \text{if} \quad (x,y) \in \Gamma, (\mu n_x + \nu n_y)(x,y) < 0 .$$

Equation (1.1) is the neutron transport equation.: the function  $u(x,y,\mu,\nu)$  represents the flux of neutrons at the point  $(x,y)$  in the angular direction  $(\mu,\nu)$ ,  $\sigma$  is the nuclear cross section and  $f$  stands for the scattering, the fission and the inhomogeneous source terms. The boundary condition (1.2) simply means that no neutrons are entering the system from outside.

In this paper, we shall be only concerned with the spatial discretization of problem (1.1), (1.2). Thus, we shall assume that the angular direction  $(\mu,\nu)$  is fixed and we shall consider the reduced problem : Given a function  $f$  defined over  $\Omega$ , find a function  $u$  defined over  $\Omega$  such that

$$(1.3) \quad \begin{cases} \mu \frac{\partial u}{\partial x} + \nu \frac{\partial u}{\partial y} + \sigma u = f \text{ in } \Omega , \\ u = 0 \text{ on } \Gamma_- , \end{cases}$$

where

$$(1.4) \quad \Gamma_- = \{ (x,y) \mid (x,y) \in \Gamma, (\mu n_x + \nu n_y)(x,y) < 0 \} .$$

This paper will be devoted to the numerical approximation of problem (1.3) by a finite element method using triangular or quadrilateral elements which has been recently introduced by Reed and Hill [17] and which appears to be very effective in practice. Other finite element methods for solving the neutron transport equation have been introduced by several authors (cf. for instance [10], [14], [15], [16]). We refer to [12] for a mathematical discussion of some of them.

An outline of the paper is as follows. In § 2, we study a discontinuous Galerkin method for ordinary differential equations using polynomials of degree  $k$ . This Galerkin method is shown to be strongly A-stable and of order  $2k + 1$ . In § 3, we introduce the finite element method as a generalization of the discontinuous Galerkin method of § 2. We prove the existence and uniqueness of the approximate solution and we give an algorithm for computing this approximate solution. In § 4, we derive general error bounds in  $L^2$ -norm. Finally, we give in § 5 a super convergence result.

Note that problem (1.3) is a simple but important example of a first order hyperbolic problem. In fact, the finite element method studied in this paper provides an effective way for numerically solving such problems. For other finite element methods for solving first order systems of partial differential equations, we refer to [11], [13].

For the sake of simplicity, we have confined ourselves to polygonal domains  $\Omega$ . It is probably an easy matter to handle general curved domains by using curved isoparametric elements and the analysis given in [5], [6].

## 2 - A DISCONTINUOUS GALERKIN METHOD FOR ORDINARY DIFFERENTIAL EQUATIONS

We begin by studying the numerical solution of the ordinary differential equation

$$(2.1) \quad \begin{cases} u'(x) = f(x, u(x)) & , \quad x \geq x_0 & , \\ u(x_0) = u_0 & , \end{cases}$$

on a finite interval  $[x_0, x_0 + a]$  by a discontinuous Galerkin method.

For continuous Galerkin methods and related collocation methods, we refer for instance to Axelsson [ 1 ] , de Boor and Swartz [ 2 ] , Hulme [ 9 ] .

Let  $x_n = x_0 + nh$  ,  $0 \leq n \leq N$  ( $Nh = a$ ) be a uniform mesh for the sake of simplicity. Then, we may approximate  $u$  on each subinterval  $[x_n, x_{n+1}]$  by a kth degree polynomial  $u_h$  . We require that  $u_h$  satisfies on each subinterval  $[x_n, x_{n+1}]$  ,  $0 \leq n \leq N-1$  :

$$(2.2) \quad \begin{cases} (u_h(x_{n+1}^-) - u_h(x_n^-))v(x_n) + \\ + \int_{x_n}^{x_{n+1}} \{u_h'(x) - f(x, u_h(x))\} v(x) dx = 0 \text{ for all } v \in P_k \end{cases}$$

with the initial condition

$$(2.3) \quad u_h(x_{0-}) = u_0 ,$$

where  $P_k$  denotes the space of all polynomials of degree  $\leq k$  . Notice that the function  $u_h$  is discontinuous in general at the mesh points  $x_n$  .

To obtain a computational form of (2.2)-(2.3), we replace the integral in (2.2) by an interpolatory quadrature formula

$$(2.4) \quad \int_{x_n}^{x_{n+1}} \varphi(x) dx = h \sum_{i=1}^{k+1} b_i \varphi(x_{n,i}) + O(h^{p+1}),$$

$$(2.5) \quad x_{n,i} = x_n + \xi_i h, \quad 1 \leq i \leq k+1, \quad \xi_1 = 0,$$

where  $b_i$  and  $\xi_i$  are the weights and abscissae for  $[0,1]$ .

Notice that  $k+1 \leq p \leq 2k+1$ . Then (2.2) becomes :

$$(2.6) \quad \begin{aligned} & (u_h(x_{n+}) - u_h(x_{n-}))v(x_n) + \\ & + h \sum_{i=1}^{k+1} b_i \{u'_h(x_{n,i}) - f(x_{n,i}, u_h(x_{n,i}))\} v(x_{n,i}) = 0 \end{aligned}$$

for all  $v \in P_k$

Let us now show that the discrete Galerkin method (2.3), (2.6) is equivalent to some implicit Runge-Kutta method. We define

$$(2.7) \quad \begin{cases} u_n = u_h(x_{n-}), \\ u_{n,1} = u_h(x_{n+}) = u_h(x_{n,1}), \\ u_{n,i} = u_h(x_{n,i}), \quad 2 \leq i \leq k+1 \end{cases}$$

We introduce the Lagrange interpolation coefficients

$$(2.8) \quad \ell_i(x) = \prod_{\substack{j=2 \\ j \neq i}}^{k+1} \frac{x - \xi_j}{\xi_i - \xi_j}, \quad 2 \leq i \leq k+1.$$

Lemma 1 The discrete Galerkin method (2.3), (2.6) is equivalent to the following implicit Runge-Kutta method

$$(2.9) \quad \begin{cases} u_{n,i} = u_n + h \sum_{j=1}^{k+1} a_{ij} f(x_{n,j}, u_{n,j}), \quad 1 \leq i \leq k+1, \\ u_{n+1} = u_n + h \sum_{j=1}^{k+1} b_j f(x_{n,j}, u_{n,j}), \end{cases}$$

where

$$(2.10) \quad \left\{ \begin{array}{l} a_{i1} = b_1, \quad 1 \leq i \leq k+1, \\ a_{ij} = \int_0^{\xi_i} l_j(x) dx - b_1 l_j(\xi_1), \quad 1 \leq i \leq k+1, \quad 2 \leq j \leq k+1 \end{array} \right.$$

Proof Let us introduce the basics  $\{v_i\}_{1 \leq i \leq k+1}$  of the space

$P_k$  defined by

$$v_i(x_{n,j}) = \delta_{ij}, \quad 1 \leq i, j \leq k+1$$

By replacing successively in (2.6)  $v$  by  $v_i$ , we find that an equivalent form of (2.6) is given by

$$(2.11) \quad \left\{ \begin{array}{l} u_h(x_{n+}) - u_h(x_{n-}) + hb_1(u_h'(x_{n,1}) - f(x_{n,1}, u_h(x_{n,1}))) = 0 \\ u_h'(x_{n,i}) - f(x_{n,i}, u_h(x_{n,i})) = 0, \quad 2 \leq i \leq k+1 \end{array} \right.$$

In the subinterval  $[x_n, x_{n+1}]$ , we have  $u_h' \in P_{k-1}$  so that

$$u_h'(x) = \sum_{j=2}^{k+1} l_j\left(\frac{x-x_n}{h}\right) u_h'(x_{n,j})$$

and by (2.11)

$$(2.12) \quad u_h'(x) = \sum_{j=2}^{k+1} l_j\left(\frac{x-x_n}{h}\right) f(x_{n,j}, u_h(x_{n,j})).$$

Taking  $x = x_n = x_{n,1}$  in (2.12), substituting this expression in the 1st equation (2.11) and using (2.7), we obtain

$$(2.13) \quad u_{n,1} = u_n + hb_1 \{f(x_{n,1}, u_{n,1}) - \sum_{j=2}^{k+1} l_j(\xi_1) f(x_{n,j}, u_{n,j})\}.$$

On the other hand, we may write for  $2 \leq i \leq k+1$

$$u_h(x_{n,i}) = u_h(x_{n,1}) + \int_{x_{n,1}}^{x_{n,i}} u'_h(x) dx$$

and by (2.7), (2.12), (2.13)

$$(2.14) \quad u_{n,i} = u_n + h \{ b_1 f(x_{n,1}, u_{n,1}) + \sum_{j=2}^{k+1} \left( \int_0^{\xi_i} l_j(x) dx - b_1 l_j(\xi_1) \right) f(x_{n,j}, u_{n,j}) \}$$

Similarly, we have

$$u_h(x_{n+1}) = u_h(x_{n,1}) + \int_{x_n}^{x_{n+1}} u'(x) dx$$

and then

$$u_{n+1} = u_n + h \{ b_1 f(x_{n,1}, u_{n,1}) + \sum_{j=2}^{k+1} \left( \int_0^1 l_j(x) dx - b_1 l_j(\xi_1) \right) f(x_{n,j}, u_{n,j}) \}$$

By noticing that

$$\int_0^1 l_j(x) dx = \sum_{i=1}^{k+1} b_i l_j(\xi_i) = b_1 l_j(\xi_1) + b_j, \quad ,$$

we get

$$(2.15) \quad u_{n+1} = u_n + h \sum_{j=1}^{k+1} b_j f(x_{n,j}, u_{n,j}) .$$

The equations (2.13) - (2.15) are identical to the equations (2.9), (2.10). We then have proved that the discrete Galerkin method leads to the one-step method (2.9), (2.10). Conversely, the Runge-Kutta method (2.9), (2.10) can be clearly viewed as a discrete Galerkin method.



Theorem 1    The discrete Galerkin method (2.3), (2.6) is a one-step method of order  $p$ .

Proof    Following Butcher [ 3 ], Crouzeix [ 7 ] , we know that the conditions

$$(2.16) \quad \sum_{j=1}^{k+1} b_j \xi_j^\ell = \frac{1}{\ell+1} , \quad 0 \leq \ell \leq p-1 ,$$

$$(2.17) \quad \sum_{j=1}^{k+1} a_{ij} \xi_j^{\ell+1} = \frac{\xi_i^{\ell+1}}{\ell+1} , \quad 0 \leq \ell \leq k-1 , \quad 1 \leq i \leq k+1$$

$$(2.18) \quad \sum_{i=1}^{k+1} b_i a_{ij} \xi_i^\ell = \frac{1}{\ell+1} b_j (1 - \xi_j^{\ell+1}) , \quad k+\ell \leq p-1 , \quad 1 \leq j \leq k+1 ,$$

are sufficient for the Runge-Kutta method (2.9) to be of order  $p$ .

Let us show that these conditions hold in the present case.

First, conditions (2.16) simply mean that the interpolatory quadrature formula (2.4) is exact for all polynomials of degree  $\leq p-1$ .

Next, consider conditions (2.17). Using (2.8) , we may write

$$x^\ell = \sum_{j=2}^{k+1} l_j(x) \xi_j^\ell , \quad 0 \leq \ell \leq k-1 ,$$

so that

$$\xi_1^\ell = \sum_{j=2}^{k+1} l_j(\xi_1) \xi_j^\ell , \quad 0 \leq \ell \leq k-1 ,$$

$$\frac{\xi_i^{\ell+1}}{\ell+1} = \sum_{j=2}^{k+1} \left( \int_0^{\xi_i} l_j(x) dx \right) \xi_j^\ell , \quad 0 \leq \ell \leq k-1 , \quad 1 \leq i \leq k+1$$

Using (2.10) , we have

$$\sum_{j=1}^{k+1} a_{ij} \xi_j^\ell = b_1 (\xi_1^\ell - \sum_{j=2}^{k+1} l_j(\xi_1) \xi_j^\ell) + \sum_{j=2}^{k+1} \left( \int_0^{\xi_i} l_j(x) dx \right) \xi_j^\ell$$

and by the previous relations

$$\sum_{j=1}^{k+1} a_{ij} \xi_j^\ell = \frac{\xi_i^{\ell+1}}{\ell+1}, \quad 0 \leq \ell \leq k-1, \quad 1 \leq i \leq k+1$$

Finally, let us show that conditions (2.18) hold . We begin by noticing that

$$(2.19) \quad \sum_{i=1}^{k+1} b_i a_{i1} \xi_i^\ell = b_1 \sum_{i=1}^{k+1} b_i \xi_i^\ell = \frac{b_1}{\ell+1}, \quad 0 \leq \ell \leq p-1.$$

On the other hand, following Crouzeix [ 7 ], we may write for all continuous function  $\varphi$

$$(2.20) \quad \int_0^1 x^\ell \left( \int_0^x \varphi(y) dy \right) dx = \frac{1}{\ell+1} \int_0^1 (1-x)^{\ell+1} \varphi(x) dx.$$

Taking  $\varphi \in P_{k-1}$  , we obtain for  $k+\ell \leq p-1$

$$\begin{aligned} \int_0^1 x^\ell \left( \int_0^x \varphi(y) dy \right) dx &= \sum_{i=1}^{k+1} b_i \xi_i^\ell \int_0^{\xi_i} \varphi(y) dy = \\ &= \sum_{i=1}^{k+1} b_i \xi_i^\ell \sum_{j=2}^{k+1} \left( \int_0^{\xi_i} l_j(y) dy \right) \varphi(\xi_j) \end{aligned}$$

and by (2.10)

$$(2.21) \quad \int_0^1 x^\ell \left( \int_0^x \varphi(y) dy \right) dx = \sum_{i,j=1}^{k+1} b_i a_{ij} \xi_i^\ell \varphi(\xi_j), \quad \varphi \in P_{k-1}, \quad k+\ell \leq p-1$$

Similarly, we get

$$(2.22) \quad \int_0^1 (1-x^{\ell+1})\varphi(x) dx = \sum_{j=1}^{k+1} b_j (1-\xi_j^{\ell+1})\varphi(\xi_j), \quad \varphi \in P_{k-1}, \quad k+\ell \leq p-1.$$

Hence, combining (2.19) - (2.22), we have for all  $\varphi \in P_{k-1}$  and for  $k+\ell \leq p-1$

$$\sum_{j=1}^{k+1} \left( \sum_{i=1}^{k+1} b_i a_{ij} \xi_i^\ell - \frac{1}{\ell+1} b_j (1-\xi_j^{\ell+1}) \right) \varphi(\xi_j) = 0$$

This implies

$$\sum_{i=1}^{k+1} b_i a_{ij} \xi_i^\ell = \frac{1}{\ell+1} b_j (1-\xi_j^{\ell+1}), \quad k+\ell \leq p-1, \quad 2 \leq j \leq k+1 \quad \square$$

For investigating the stability properties of the one-step method (2.9), we consider the differential equation

$$(2.23) \quad u' = \lambda u$$

where  $\lambda$  is a complex constant with  $\operatorname{Re}(\lambda) < 0$ .

Lemma 2. Applied to the differential equation (2.23), the one-step method (2.9), (2.10) gives

$$(2.24) \quad u_{n+1} = R(\lambda h) u_n$$

where  $R(z) = \frac{P(z)}{Q(z)}$  is the quotient of two polynomials  $P(z)$  and  $Q(z)$  of degree  $\leq k$  and  $k+1$  respectively.

Proof Applied to (2.23), the one-step method (2.9) becomes (2.25)

$$(2.25) \quad u_{n,i} = u_n + \lambda h \sum_{j=1}^{k+1} a_{ij} u_{n,j}, \quad 1 \leq i \leq k+1,$$

$$(2.26) \quad u_{n+1} = u_n + \lambda h \sum_{j=1}^{k+1} b_j u_{n,j}.$$

Using obvious notations, we may write equations (2.25) in the form

$$(I - \lambda h [a_{ij}]) [u_{n,i}] = u_n [1]$$

where the identity matrix  $I$  and  $[a_{ij}]$  are  $(k+1) \times (k+1)$ -matrices.

Since  $a_{i1} = b_1$ ,  $1 \leq i \leq k+1$ , we get from Cramer's rule

$$u_{n,i} = \frac{P_i(\lambda h)}{Q(\lambda h)} u_n, \quad 1 \leq i \leq k+1$$

where  $P_1(z)$  is a polynomial of degree  $k$  whose leading coefficient is  $b_1^{-1} \det(a_{ij})$ ,  $P_i(z)$ ,  $2 \leq i \leq k+1$ , are polynomials of degree  $\leq k-1$  and where  $Q(z)$  is a polynomial of degree  $k+1$  whose leading coefficient is  $\det(a_{ij})$ .

Using (2.26), we obtain

$$u_{n+1} = \frac{P(\lambda h)}{Q(\lambda h)} u_n$$

where

$$P(z) = Q(z) - z \sum_{j=1}^{k+1} P_j(z).$$

Clearly, in  $P(z)$ , the coefficient of  $z^{k+1}$  vanishes. The lemma is then proved.  $\square$

Let us now recall the following definition. A one-step method is strongly A-stable if

$$(2.27) \quad \begin{cases} |R(z)| < 1 & \text{for } \operatorname{Re}(z) < 0 \\ |R(z)| \rightarrow 0 & \text{as } \operatorname{Re}(z) \rightarrow -\infty \end{cases}$$

Theorem 2     The Galerkin method (2.2) , (2.3) is a strongly A-stable one-step method of order  $2k+1$  .

Proof     Consider first the discrete Galerkin method (2.3) , (2.6) associated with the Gauss-Radau abscissae  $\xi_i$  ,  $1 \leq i \leq k+1$  ( $\xi_1 = 0$ ). Then, we have  $p = 2k+1$  in (2.4). By Theorem 1 , this discrete Galerkin method is a one-step method of order  $2k+1$  so that

$$R(z) = \exp(z) + O(z^{2k+2}) .$$

Moreover, by lemma 2 ,  $R(z)$  is the quotient of two polynomials  $P(z)$  and  $Q(z)$  of degree  $\leq k$  and  $\leq k+1$  respectively. Then, necessarily ,  $R(z)$  is the subdiagonal  $(k+1,k)$  Padé rational approximation of  $\exp(z)$ . Using a result of Axelsson [ 1 ] , we know that such a Padé approximation satisfies conditions (2.27). Hence, the discrete Galerkin method (2.3) , (2.6) associated with the Gauss-Radau abscissae is a strongly A-stable one-step method of order  $2k+1$  .

Now, it is a simple but lengthy matter to prove that the Galerkin method (2.2) , (2.3) and the Gauss-Radau discrete Galerkin method (2.3) (2.6) are one-step methods of the same order  $2k+1$  . Moreover, these two methods coincide when applied to the differential equation (2.23). This completes the proof of the theorem.      $\square$

### 3 - A FINITE ELEMENT METHOD FOR THE NEUTRON TRANSPORT EQUATION

Consider now our neutron transport problem (1.3). First, we need some notations. Let us denote by  $L^2(\Omega)$  the space of real-valued functions  $v$  which are square integrable over  $\Omega$ .

We provide  $L^2(\Omega)$  with the usual norm

$$(3.1) \quad \|v\|_{0,\Omega} = \left( \int_{\Omega} |v(x)|^2 dx \right)^{1/2}.$$

Given any integer  $m \geq 0$ , let

$$(3.2) \quad H^m(\Omega) = \{v | v \in L^2(\Omega), \partial^\alpha v \in L^2(\Omega), |\alpha| \leq m\}$$

be the usual Sobolev space provided with the norm

$$(3.3) \quad \|v\|_{m,\Omega} = \left( \sum_{|\alpha| \leq m} \|\partial^\alpha v\|_{0,\Omega}^2 \right)^{1/2}$$

In (3.2), (3.3),  $\alpha = (\alpha_1, \alpha_2) \in \mathbb{N}^2$  is a multiindex,  $|\alpha| = \alpha_1 + \alpha_2$

and

$$\partial^\alpha = \left( \frac{\partial}{\partial x_1} \right)^{\alpha_1} \left( \frac{\partial}{\partial x_2} \right)^{\alpha_2}$$

We shall also use the following semi-norm

$$(3.4) \quad |v|_{m,\Omega} = \left( \sum_{|\alpha|=m} \|\partial^\alpha v\|_{0,\Omega}^2 \right)^{1/2}$$

Let us introduce the operator

$$(3.5) \quad A = \mu \frac{\partial}{\partial x} + \nu \frac{\partial}{\partial y} + \sigma$$

and the space

$$(3.6) \quad D(A) = \{v | v \in L^2(\Omega), \mu \frac{\partial v}{\partial x} + \nu \frac{\partial v}{\partial y} \in L^2(\Omega)\}$$

Then, as a consequence of [8 ], we have the following result.

Theorem 3 . Assume that  $\sigma \in L^\infty(\Omega)$  and  $f \in L^2(\Omega)$  . Then,  
problem (1.3) has a unique strong solution  $u \in D(A)$  .

Using the change of unknown function

$$u = \exp\left(\lambda\left(\frac{x}{\mu} + \frac{y}{\nu}\right)\right)w ,$$

equation (1.3) becomes

$$\mu \frac{\partial w}{\partial x} + \frac{\partial w}{\partial y} + (\sigma + \lambda)w = \exp\left(-\lambda\left(\frac{x}{\mu} + \frac{y}{\nu}\right)\right)f .$$

Thus, by eventually changing  $\sigma(x,y)$  into  $\sigma(x,y) + \lambda$ , we can confine ourselves to the case where  $\sigma$  is positive. More precisely, we shall assume in all the sequel that

$$(3.7) \quad M \geq \sigma(x,y) \geq \alpha > 0 \quad \text{a.e. in } \Omega .$$

Let us now generalize the one-dimensional discontinuous Galerkin method of § 2 to our two-dimensional neutron transport problem. For the sake of simplicity, we shall assume in the following that  $\Omega$  is a polygon. In order to approximate problem (1.3) , we first construct a triangulation  $\mathcal{T}_h$  of  $\bar{\Omega}$  with triangles and convex quadrilaterals  $K$  with diameters  $\leq h$  . With any  $K \in \mathcal{T}_h$  , we associate a finite-dimensional space  $P_K$  of real-valued functions defined over  $K$  such that

$$(3.8) \quad P_K \subset H^1(K) .$$

We then consider the finite-dimensional space

$$(3.9) \quad V_h = \{v \mid v \in L^2(\Omega) , v|_K \in P_K \text{ for all } K \in \mathcal{T}_h\} .$$

It is worthwhile to notice that in general a function  $v \in V_h$  does not satisfy any continuity requirement at the interelement boundaries.

Let  $K \in \mathcal{C}_h$  and let  $\partial K$  be the boundary of  $K$ . We set :

$$(3.10) \left\{ \begin{array}{l} \partial_- K = \{(x,y) \mid (x,y) \in \partial K, (\mu n_x + \nu n_y)(x,y) < 0\} , \\ \partial_+ K = \{(x,y) \mid (x,y) \in \partial K, (\mu n_x + \nu n_y)(x,y) > 0\} , \end{array} \right.$$

where  $n_x$  and  $n_y$  are the components of the outward unit vector normal to the boundary  $\partial K$ .

Then, the finite element approximation of problem (1.3) that we shall consider here can be stated as follows. Find a function  $u_h \in V_h$  such that for all  $K \in \mathcal{C}_h$

$$(3.11) \quad \int_{\partial_- K} (\mu n_x + \nu n_y) (u_h - \xi_h) v \, ds + \int_K (A u_h - f) v \, dx dy = 0 \quad \forall v \in P_K$$

where

$$(3.12) \quad \xi_h = \begin{cases} 0 & \text{on } \partial_- K \cap \Gamma_- \\ \text{outward trace of } u_h & \text{on } \partial_- K - (\partial_- K \cap \Gamma_-) \end{cases}$$

Clearly, this method appears to be the direct generalization of the discontinuous Galerkin method (2.2) , (2.3).

Before proving existence and uniqueness of the solution  $u_h \in V_h$  we shall show that there exists an ordering of the elements of  $\mathcal{C}_h$  well suited for numerically solving equations (3.11) , (3.12).

Lemma 3 There exists an ordering  $K_1, K_2, \dots, K_I$  of the elements of  $\mathcal{C}_h$  such that, for all  $i = 1, \dots, I$ , each side of  $\partial_- K_i$  is either a subset of  $\Gamma_-$  or a subset of  $\partial_+ K_j$  for some  $j < i$ .



Proof Let us introduce first some notations. We shall say that  $K$  is a boundary element if at least one side of  $\partial K$  is a subset of  $\Gamma$  and that  $K$  is a semi-boundary element if one and only one vertex of  $K$  belongs to  $\Gamma$ . Let us consider  $\Gamma_-$  and let us number clockwise the corresponding

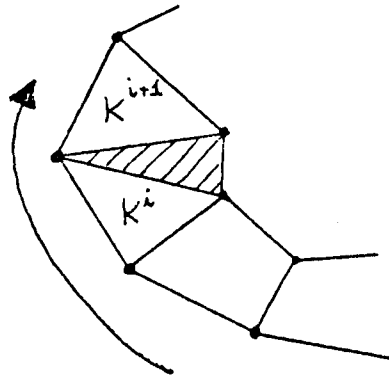


Fig. 1

boundary elements  $K^1, K^2, \dots, K^s$ . Two consecutive boundary elements  $K^i$  and  $K^{i+1}$  can have a common side or not. In the latter case (cf. Fig. 1), there exists at least one semi boundary element located between  $K^i$  and  $K^{i+1}$ . Then, we shall say that a side of  $K^i$  (resp.  $K^{i+1}$ ) is semi-common with  $K^{i+1}$  (resp.  $K^i$ ) if it is a subset of the union of the semi-boundary elements located between  $K^i$  and  $K^{i+1}$ .

Next, we show that there exists at least one boundary element  $K$  such that  $\partial_K \subset \Gamma_-$ . To prove this, let us assume on the contrary that  $\partial_{K^i} \not\subset \Gamma_-$  for all  $i = 1, \dots, s$  and let us show that this assumption leads to a contradiction.

Consider the first boundary element  $K^1$  and use the notations of Fig. 2.

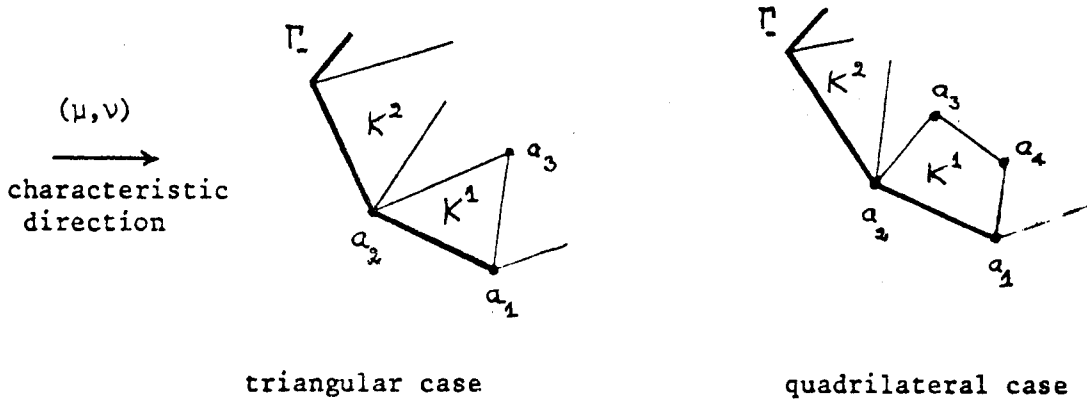


Fig. 2

In the triangular case (resp. in the quadrilateral case), the side  $[a_1, a_3]$  (resp.  $[a_1, a_4]$ ) of  $K^1$  is a subset of  $\partial_+ K^1$ . Otherwise,  $K^1$  would not be the first boundary element of  $\Gamma_-$ . Then, the side  $[a_2, a_3]$  of  $K^1$  which is common or semi-common with  $K^2$  belongs to  $\partial_- K^1$ . Otherwise, we should get  $\partial_- K^1 = [a_1, a_2] \subset \Gamma_-$  which is excluded. Therefore, the side of  $K^2$  which is common or semi-common with  $K^1$  belongs to  $\partial_+ K^2$ . More generally, we get for every  $i = 1, \dots, s-1$  the following property : the side of  $K^i$  which is common or semi common with  $K^{i+1}$  is a subset of  $\partial_- K^i$  and therefore the side of  $K^{i+1}$  which is common or semi-common with  $K^i$  is a subset of  $\partial_+ K^{i+1}$ . Now consider the last boundary element  $K^s$  and use the notations of Fig. 3

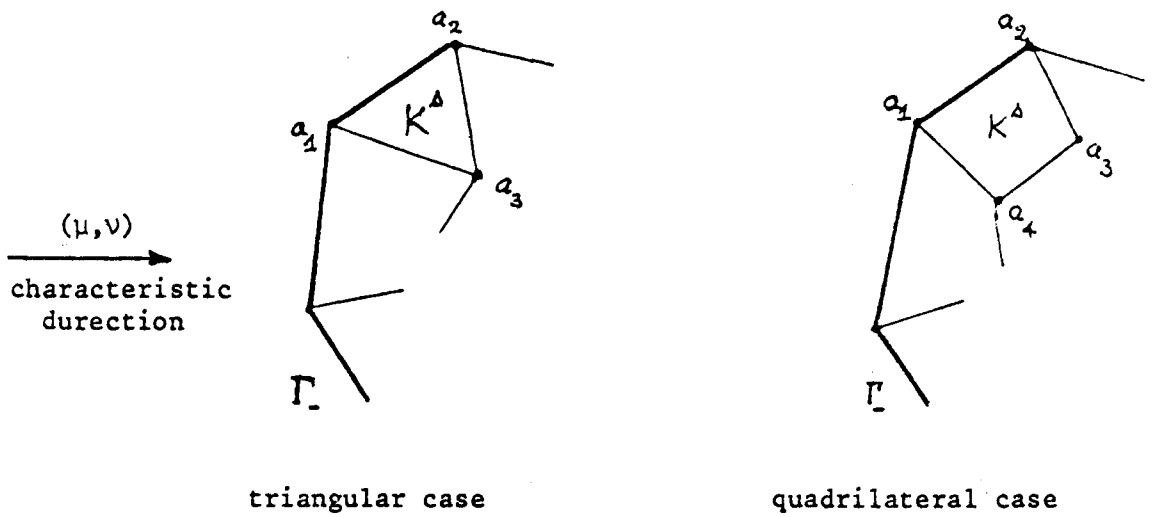


Fig. 3

In the triangular case (resp. in the quadrilateral case), the side  $[a_1, a_3]$  (resp.  $[a_1, a_4]$ ) of  $K^s$  is a subset of  $\partial_+ K^s$ . Moreover, the side  $[a_2, a_3]$  is a subset of  $\partial_+ K^s$ . Otherwise,  $K^s$  would not be the last boundary element of  $\Gamma_-$ . Thus, we get  $\partial_- K^s = [a_1, a_2] \subset \Gamma_-$  which has been excluded. The existence of a boundary element  $K$  such that  $\partial_- K \subset \Gamma_-$  is then proved.

Now, choose for  $K_1$  a boundary element of  $\Gamma_-$  such that  $\partial_- K_1 \subset \Gamma_-$  and define  $\Omega_1 = \Omega - \Omega \cap K_1$ ,  $\Gamma_{1-} = \partial_- \Omega_1$ . Note that each side of  $\Gamma_{1-}$  is either a subset of  $\Gamma_-$  or a subset of  $\partial_+ K_1$ . By the previous argument, there exist a boundary element  $K_2$  of  $\Gamma_{1-}$  such that  $\partial_- K_2 \subset \Gamma_{1-}$ , etc ... . Repeating this process, we take into account all the elements of  $\mathcal{G}_h$  and we obtain an ordering  $K_1, K_2, \dots, K_I$  of the elements of  $\mathcal{G}_h$  such that the desired property holds.  $\square$

This proof suggests an ordering algorithm for the elements of

$\mathcal{T}_h$  which is effectively used in practise.

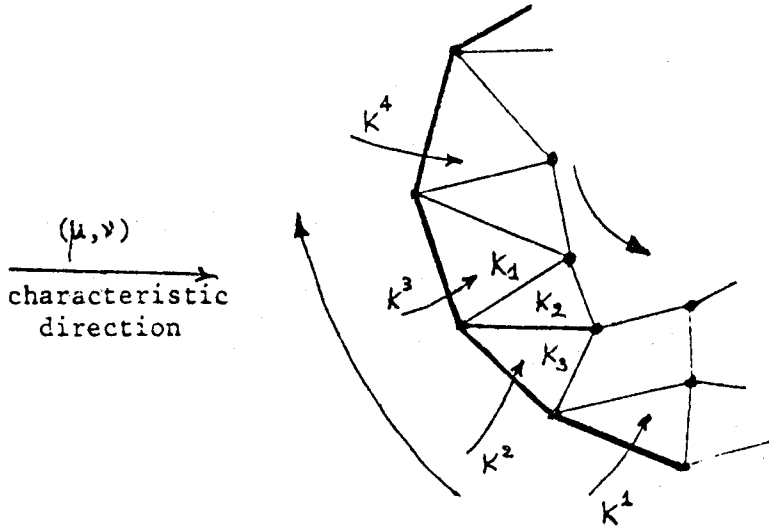


Fig. 4

Consider the sequence  $K^1, K^2, \dots, K^S$  of boundary elements of  $\Gamma_-$ . For  $K_1$  we choose the first element  $K$  of this sequence which satisfies  $\partial_- K \subset \Gamma_-$ . Let  $K^p$  be this element ( $p=3$  in Fig.4). From  $K_1 = K^p$ , we then number counterclockwise  $K_1, K_2, \dots, K_r$  the boundary and semi boundary elements located between  $K^p$  and  $K^1$  which satisfy the following condition : for all  $i = 1, \dots, r$ , each side of  $\partial_- K_i$  is either a subset of  $\Gamma_-$  or a subset of  $\partial_+ K_j$  for some  $j < i$  ( $r=3$  in Fig.4). Next, we replace the set  $\Omega$  by  $\Omega \setminus \Omega \cap (\bigcup_{i=1}^r K_i)$  and we repeat the process ...

We are now able to prove

Theorem 4 Assume that  $f \in L^2(\Omega)$  and that condition (3.7) holds.  
Then, there exists a unique function  $u_h \in V_h$  which satisfies  
equations (3.11) and (3.12) for all  $K \in \mathcal{T}_h$ .

Proof Clearly, the finite element method (3.11), (3.12) is equivalent to a  $N \times N$  linear system of equations with  $N = \dim V_h$ . Then, it is sufficient to prove the uniqueness of the solution  $u_h$ . Thus, let us assume that  $f = 0$  and let us show that necessarily  $u_h = 0$ . Let  $K_1, K_2, \dots, K_I$  be an ordering of the elements  $K \in \mathcal{C}_h$  such that the condition of Lemma 3 holds. If  $u_h = 0$  in  $K_1 \cup K_2 \cup \dots \cup K_{i-1}$ , then  $\xi_h = 0$  on  $\partial_- K_i$  and equation (3.11) becomes in  $K = K_i$

$$-\int_{\partial_- K_i} (\mu n_x + \nu n_y) u_h v \, ds + \int_{K_i} (A u_h) v \, dx dy = 0 \quad \forall v \in P_{K_i}$$

Taking  $v = u_h$  and using Green's formula

$$\int_{K_i} \left( \mu \frac{\partial u_h}{\partial x} + \nu \frac{\partial u_h}{\partial y} \right) u_h \, dx dy = \frac{1}{2} \int_{\partial K_i} (\mu n_x + \nu n_y) u_h^2 \, ds,$$

we get

$$\int_{\partial_+ K_i} (\mu n_x + \nu n_y) u_h^2 \, ds - \int_{\partial_- K_i} (\mu n_x + \nu n_y) u_h^2 \, ds + \int_{K_i} \sigma u_h^2 \, dx dy = 0$$

Using (3.7) and (3.10), we obtain  $u_h = 0$  in  $K_i$ . Therefore, using an inductive argument, we get  $u_h = 0$  in  $\Omega$ .  $\square$

In practice, the computation of the approximate solution  $u_h \in V_h$  goes along the following lines :

(i) Find an ordering  $K_1, K_2, \dots, K_I$  of the elements  $K \in \mathcal{C}_h$  which satisfies the condition of Lemma 3, for instance by using the previous algorithm ;

(ii) Compute successively  $u_h$  in  $K_1, K_2, \dots, K_I$ . The computation of  $u_h$  in each  $K_i$  has a local character and involves the numerical solution of a  $d_i \times d_i$  linear system where  $d_i = \dim P_{K_i}$ .

In other words, by using an ordering of  $\mathcal{T}_h$  such that the condition of Lemma 3 holds, the  $N \times N$  matrix of the approximate problem becomes block triangular and the  $i$ th diagonal block is a  $d_i \times d_i$  matrix associated with the  $i$ th element  $K_i$ .

Note that, in many practical problems, the geometry of  $\Omega$  and the triangulation  $\mathcal{T}_h$  are so simple that step (i) becomes obvious.

#### 4 - GENERAL ERROR BOUNDS

Let us now derive some estimates for the error  $u_h - u$  when the solution  $u$  of problem (1.3) is smooth enough. We begin with

Lemma 4 For any  $K \in \mathcal{E}_h$ , any  $v \in P_K$  and any function  $\eta \in L^2(\partial_-K)$ , we have the estimate :

$$\begin{aligned}
 (4.1) \quad & \left. \begin{aligned}
 & \frac{1}{2} \int_{\partial_+K} (\mu n_x + \nu n_y) (u_h - v)^2 ds + \frac{1}{2} \int_{\partial_-K} (\mu n_x + \nu n_y) (\xi_h - \eta)^2 ds - \\
 & - \frac{1}{2} \int_{\partial_-K} (\mu n_x + \nu n_y) ((u_h - v) - (\xi_h - \eta))^2 ds + \int_K \sigma (u_h - v)^2 dx dy = \\
 & = \int_{\partial_+K} (\mu n_x + \nu n_y) (u - v) (u_h - v) ds + \int_{\partial_-K} (\mu n_x + \nu n_y) (u - \eta) (u_h - v) ds \\
 & + \int_K (u - v) A^* (u_h - v) dx dy
 \end{aligned} \right\}
 \end{aligned}$$

where  $A^*$  is the formal adjoint of the operator  $A$ , i.e.

$$(4.2) \quad A^* = -\mu \frac{\partial}{\partial x} - \nu \frac{\partial}{\partial y} + \sigma.$$

Proof Given  $p \in P_K$  and  $\eta \in L^2(\partial_-K)$ , we set :

$$(4.3) \quad w = u_h - v \in P_K, \quad \zeta = \xi_h - \eta.$$

Consider the expression

$$(4.4) \quad X_h = - \int_{\partial_-K} (\mu n_x + \nu n_y) (w - \zeta) w ds + \int_K (Aw) w dx dy$$

First, using Green's formula, we obtain

$$X_h = \frac{1}{2} \int_{\partial K} (\mu n_x + \nu n_y) w^2 ds - \int_{\partial_- K} (\mu n_x + \nu n_y) (w-\zeta) w ds + \int \sigma w^2 dx dy .$$

Since

$$(w-\zeta)w = \frac{1}{2}(w^2 - \zeta^2 + (w-\zeta)^2) ,$$

we get

$$(4.5) \quad \left\{ \begin{array}{l} X_h = \frac{1}{2} \int_{\partial_+ K} (\mu n_x + \nu n_y) w^2 ds + \frac{1}{2} \int_{\partial_- K} (\mu n_x + \nu n_y) \zeta^2 ds - \\ - \frac{1}{2} \int_{\partial_- K} (\mu n_x + \nu n_y) (w-\zeta)^2 ds + \int_K \sigma w^2 dx dy \end{array} \right.$$

On the other hand, using (3.11) , we obtain

$$X_h = \int_{\partial_- K} (\mu n_x + \nu n_y) (v-\eta) w ds + \int_K (f-Av) w dx dy$$

and therefore

$$X_h = \int_{\partial_- K} (\mu n_x + \nu n_y) (v-\eta) w ds + \int_K A(u-v) w dx dy$$

Since  $u \in D(A)$  , we may write

$$\int_K A(u-v) w dx dy = \int_K (u-v) A^* w dx dy + \int_{\partial K} (\mu n_x + \nu n_y) (u-v) w ds$$

so that

$$(4.6) \quad \left\{ \begin{array}{l} X_h = \int_{\partial_+ K} (\mu n_x + \nu n_y) (u-v) w ds + \int_{\partial_- K} (\mu n_x + \nu n_y) (u-\eta) w ds + \\ + \int_K (u-v) A^* w dx dy . \end{array} \right.$$



By combining (4.3) , (4.5) and (4.5) , we get the desired estimate.  $\square$

In order to get explicit error bounds, we need to define more precisely the finite-dimensional spaces  $P_K$  . Let  $K$  be an element of  $\mathcal{C}_h$  . If  $K$  is a triangle, there exists an affine invertible mapping  $F_K$  which maps a reference triangle  $\hat{K}$  onto  $K$  ( $\hat{K}$  is usually chosen as a unit isosceles rectangular triangle). If  $K$  is a non-degenerate convex quadrilateral, there exists a bi-affine invertible mapping  $F_K$  which maps the reference element  $\hat{K} = [-1, +1]^2$  onto  $K$  . Note that this mapping  $F_K$  becomes affine when  $K$  is a parallelogram.

In both cases, let  $\hat{P} \in H^1(\hat{K})$  be a finite-dimensional space of real-valued functions defined over the reference element  $\hat{K}$  . We shall always assume in the following that

$$(4.7) \quad P_K = \{p \mid p = \hat{p} \circ F_K^{-1}, \hat{p} \in \hat{P}\} .$$

We shall make a constant use of the one-to-one correspondence

$$\hat{v} \rightarrow v = \hat{v} \circ F_K^{-1} \quad , \quad v \rightarrow \hat{v} = v \circ F_K$$

between the functions  $\hat{v}$  defined over  $\hat{K}$  and the functions  $v$  defined over  $K$  .

For any integer  $m \geq 0$  , let  $P_m$  denote the space of all polynomials of degree  $\leq m$  in the two variables  $x, y$  and let  $Q_m$  denote the space of all polynomials of the form

$$p(x, y) = \sum_{i, j=0}^m c_{ij} x^i y^j \quad , \quad c_{ij} \in \mathbb{R} .$$

We shall need

Hypothesis H.1.      There exists an integer  $k \geq 0$  such that :

$$(4.8) \quad P_k \subset \hat{P} \quad \text{if } \hat{K} \text{ is the reference triangle ,}$$

$$(4.9) \quad Q_k \subset \hat{P} \quad \text{if } \hat{K} \text{ is the reference quadrilateral } [-1, +1]^2 .$$

Let us now introduce the following geometrical parameters :

$$(4.10) \quad \left\{ \begin{array}{l} h(K) = \text{diameter of } K \\ \rho(K) = \sup \{ \text{diameter of the circles contained in } K \} , \\ \theta_i(K) , 1 \leq i \leq 4 = \text{angles of } K \text{ if } K \text{ is a quadrilateral.} \end{array} \right.$$

Hypothesis H.2.      There exists a constant  $\sigma > 1$  independent of  $h$  such that

$$(4.11) \quad h(K) \leq \sigma \rho(K) \quad \text{for all } K \in \mathcal{C}_h .$$

Moreover, there exists a constant  $\gamma$  independent of  $h$  with  $0 < \gamma < 1$  such that

$$(4.12) \quad \max_{1 \leq i \leq 4} |\cos \theta_i(K)| < \gamma \quad \text{for all quadrilateral } K \in \mathcal{C}_h .$$

Given a reference element  $\hat{K}$ , we define  $\hat{\Pi}$  to be the orthogonal projection operator in  $L^2(\hat{K})$  upon  $\hat{P}$ . For any  $K \in \mathcal{C}_h$ , we define  $\Pi_K \in \mathcal{L}(L^2(K); P_K)$  by

$$(4.13) \quad \widehat{\Pi_K v} = \hat{\Pi} \hat{v} \quad \text{for all } v \in L^2(K) .$$

Then, for any  $v \in L^2(\Omega)$ , we define  $\Pi_h v$  to be the function in  $V_h$  such that

$$(4.14) \quad \Pi_h v|_K = \Pi_K v \quad \text{for all } K \in \mathcal{C}_h .$$

Let us now state some standard results which can be easily proved by using the techniques of Ciarlet & Raviart [4], [5].

Lemma 5 Assume that hypothesis H.2 holds. Then, there exists a constant  $c > 0$  independent of  $K \in \mathcal{C}_h$  such that for all  $p \in P_K$

$$(4.15) \quad |p|_{1,K} \leq C(h(K))^{-1} \|p\|_{0,K}$$

$$(4.16) \quad \|p\|_{0,K'} \leq C(h(K))^{-1/2} \|p\|_{0,K},$$

where  $K'$  is any side of  $K$  and  $\|p\|_{0,K'} = (\int_{K'} |p|^2 ds)^{1/2}$ .

Lemma 6 Assume that hypotheses H.1, H.2 and (4.13) hold. Then, there exists a constant  $c > 0$  independent of  $K \in \mathcal{C}_h$  such that for all  $v \in H^{k+1}(K)$

$$(4.17) \quad |v - \Pi_K v|_{m,K} \leq C(h(K))^{k+1-m} \|v\|_{k+1,K}, \quad m = 0, 1,$$

$$(4.18) \quad \|v - \Pi_K v\|_{0,K'} \leq C(h(K))^{k+1/2} \|v\|_{k+1,K}$$

where  $K'$  is any side of  $K$ .

Let  $K_1, K_2, \dots, K_I$  be a fixed ordering of the elements of  $\mathcal{C}_h$  which satisfies the condition of Lemma 3. For all  $i = 1, \dots, I$ , we set

$$(4.19) \quad \Omega_i = \overset{\circ}{\bigcup_{j=1}^i K_j}$$

and we define  $\partial_+ \Omega_i$  and  $\partial_- \Omega_i$  in the usual way. Note that  $\partial_- \Omega_i \subset \Gamma_-$ .

Theorem 5    Assume that Hypotheses H.1 and H.2 hold. Assume in addition that the solution  $u$  of problem (1.3) belongs to  $H^{k+1}(\Omega)$ . Then, there exists a constant  $C > 0$  independent of  $h$  such that for all  $i = 1, \dots, I$

$$(4.20) \quad \|u_h - u\|_{0, \Omega_i} \leq Ch^k \|u\|_{k+1, \Omega_i}$$

$$(4.21) \quad \left( \int_{\partial_+ \Omega_i} (\mu n_x + \nu n_y) (u_h - u)^2 ds \right)^{1/2} \leq Ch^k \|u\|_{k+1, \Omega_i}$$

$$(4.22) \quad \left( - \sum_{j=1}^i \int_{\partial_- K_j} (\mu n_x + \nu n_y) (u_h - \xi_h)^2 ds \right)^{1/2} \leq Ch^k \|u\|_{k+1, \Omega_i}$$

Proof    For any  $K \in \mathcal{C}_h$ , we define

$$(4.23) \quad \eta_h = \begin{cases} 0 & \text{on } \partial_- K \cap \Gamma_- \\ \text{outward trace of } \Pi_h u & \text{on } \partial_- K - (\partial_- K \cap \Gamma_-) \end{cases}$$

We start from equation (4.1) with  $v = \Pi_h u$ ,  $\eta = \eta_h$  and we estimate the corresponding right hand side member. First, we have

$$\left| \int_K (u - \Pi_h u) A^* (u_h - \Pi_h u) dx dy \right| \leq c_1 \|u - \Pi_h u\|_{0, K} \|u_h - \Pi_h u\|_{1, K} \quad (1)$$

and by (4.15), (4.17)

$$(4.24) \quad \left| \int_K (u - \Pi_h u) A^* (u_h - \Pi_h u) dx dy \right| \leq c_2 h^k \|u\|_{k+1, K} \|u_h - \Pi_h u\|_{0, K}$$

Next, using (4.16) and (4.18), we obtain

$$(4.25) \quad \left| \int_{\partial_+ K} (\mu n_x + \nu n_y) (u - \Pi_h u) (u_h - \Pi_h u) ds \right| \leq c_3 h^k \|u\|_{k+1, K} \|u_h - \Pi_h u\|_{0, K}$$

---

(1) In all of the sequel, we shall denote by  $c_i$  various constants independent of  $h$ .

Similarly, we get

$$(4.26) \quad \left| \int_{\partial_- K} (\mu n_x + \nu n_y) (u - \eta_h) (u_h - \Pi_h u) ds \right| \leq c_4 h^k \|u\|_{k+1, \mathcal{D}_K} \|u_h - \Pi_h u\|_{0, K}$$

where  $\mathcal{D}_K$  is the union of the elements of  $\mathcal{E}_h$  which have a side contained in  $\partial_- K$ .

Thus, combining (4.1) with  $v = \Pi_h u$ ,  $\eta = \eta_h$ , (4.24), (4.25),

(4.26) and using (3.7), we obtain

$$\left\{ \begin{array}{l} \frac{1}{2} \int_{\partial_+ K} (\mu n_x + \nu n_y) (u_h - \Pi_h u)^2 ds - \\ - \frac{1}{2} \int_{\partial_- K} (\mu n_x + \nu n_y) ((u_h - \Pi_h u) - (\xi_h - \eta_h))^2 ds + \alpha \|u_h - \Pi_h u\|_{0, K}^2 \leq \\ \leq - \frac{1}{2} \int_{\partial_- K} (\mu n_x + \nu n_y) (\xi_h - \eta_h)^2 ds + c_5 h^k \|u\|_{k+1, K \cup \mathcal{D}_K} \|u_h - \Pi_h u\|_{0, K} \end{array} \right.$$

Summing over all the elements  $K_j$ ,  $1 \leq j \leq i$ , and using (3.12),

(4.23) we get

$$(4.27) \quad \left\{ \begin{array}{l} \frac{1}{2} \int_{\partial_+ \Omega_i} (\mu n_x + \nu n_y) (u_h - \Pi_h u)^2 ds - \\ - \frac{1}{2} \sum_{j=1}^i \int_{\partial_- K_j} (\mu n_x + \nu n_y) ((u_h - \Pi_h u) - (\xi_h - \eta_h))^2 ds + \alpha \|u_h - \Pi_h u\|_{0, \Omega_i}^2 \leq \\ \leq c_6 h^k \|u\|_{k+1, \Omega_i} \|u_h - \Pi_h u\|_{0, \Omega_i} \end{array} \right.$$

From (4.17) and (4.27), we deduce :

$$\begin{aligned} \|u_h - u\|_{0, \Omega_i} &\leq \|u_h - \Pi_h u\|_{0, \Omega_i} + \|\Pi_h u - u\|_{0, \Omega_i} \leq \\ &\leq \frac{c_6}{\alpha} h^k \|u\|_{k+1, \Omega_i} + c_7 h^{k+1} \|u\|_{k+1, \Omega_i} \end{aligned}$$

so that (4.20) holds.

Next, we have by (4.27)

$$\left( \int_{\partial_+ \Omega_i} (\mu n_x + \nu n_y) (u_h - \Pi_h u)^2 ds \right)^{1/2} \leq c_8 h^k \|u\|_{k+1, \Omega_i}$$

and by (4.18)

$$\left( \int_{\partial_+ \Omega_i} (\mu n_x + \nu n_y) (\Pi_h u - u)^2 ds \right)^{1/2} \leq c_9 h^{k+1/2} \|u\|_{k+1, \Omega_i}$$

This proves inequality (4.21).

Similarly, we have by (4.27)

$$\left( - \sum_{j=1}^i \int_{\partial_- K_j} (\mu n_x + \nu n_y) ((u_h - \Pi_h u) - (\xi_h - \eta_h))^2 ds \right)^{1/2} \leq c_{10} h^k \|u\|_{k+1, \Omega_i}$$

and by (4.18)

$$\left( - \sum_{j=1}^i \int_{\partial_- K_j} (\mu n_x + \nu n_y) (\Pi_h u - u)^2 ds \right)^{1/2} \leq c_{11} h^{k+1/2} \|u\|_{k+1, \Omega_i},$$

$$\left( - \sum_{j=1}^i \int_{\partial_- K_j} (\mu n_x + \nu n_y) (\eta_h - u)^2 ds \right)^{1/2} \leq c_{12} h^{k+1/2} \|u\|_{k+1, \Omega_i}$$

This implies inequality (4.22).  $\square$

5 - A SUPERCONVERGENCE RESULT

Let us notice that the error estimates of theorem 5 are not optimal in the exponent of the parameter  $h$ . In fact, numerical calculations have shown that these error bounds could not be improved in general. However, the one-dimensional results of § 2 clearly indicate that better estimates must hold in some special cases. Indeed, we shall prove in this § that the rate of convergence of our finite element method is  $O(h^{k+1})$  when all the elements  $K \in \mathcal{C}_h$  are rectangles and when  $\hat{P} = Q_k$ . In all the sequel, we shall confine ourselves to this particular case.

On the interval  $[-1, +1]$ , let  $-1 < \theta_1 < \theta_2 < \dots < \theta_{k+1} = 1$  denote the  $(k+1)$  Gauss-Radau quadrature abscissae. In the reference square  $\hat{K} = [-1, +1]^2$ , we consider the points  $\hat{a}_{ij}$  with coordinates  $(\theta_i, \theta_j)$ ,  $1 \leq i, j \leq k+1$ .

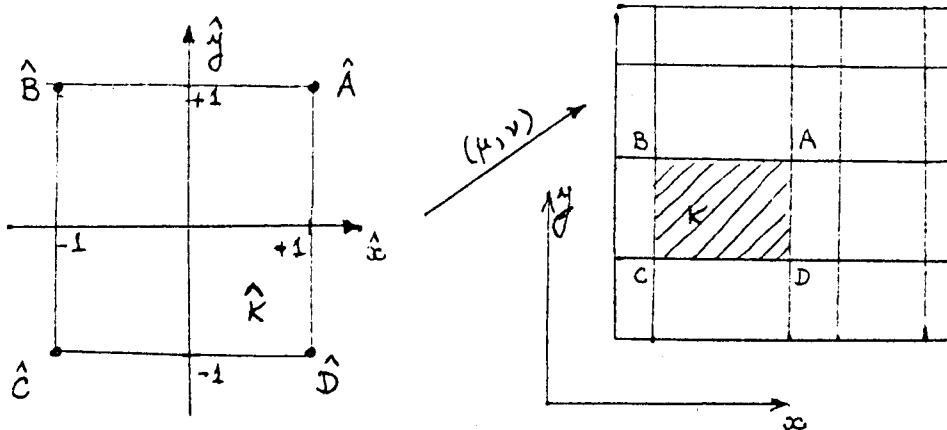


Fig. 5.

Just for convenience, we shall assume that the sides of the rectangles  $K \in \mathcal{T}_h$  are parallel to the  $(x,y)$  axes and that the coefficients  $\mu, \nu$  are  $> 0$ . Given a rectangle  $K$  with vertices  $A, B, C, D$  as in Fig. 5, we denote by  $F_K$  the affine invertible mapping such that  $A = F_K(\hat{A}), \dots, D = F_K(\hat{D})$ . Given a function  $v \in C^0(K)$ , we define  $r_K v$  as the unique polynomial of  $Q_k$  which interpolates  $v$  at the points  $a_{ij} = F_K(\hat{a}_{ij})$ ,  $1 \leq i, j \leq k+1$ . Then, for any  $v \in C^0(\bar{\Omega})$ , we define  $r_h v$  to be the function in  $V_h$  such that

$$(5.1) \quad r_h v|_K = r_K v \quad \text{for all } K \in \mathcal{T}_h.$$

We provide  $L^\infty(\Omega)$  with the following norm

$$\|v\|_{0, \infty, \Omega} = \sup \{ |v(x)| ; x \in \Omega \}$$

Given any integer  $m \geq 0$ , let

$$W^{m, \infty}(\Omega) = \{ v | v \in L^\infty(\Omega), \partial^\alpha v \in L^\infty(\Omega), |\alpha| \leq m \}$$

be the Sobolev space provided with the norm

$$\|v\|_{m, \infty, \Omega} = \max \{ \|\partial^\alpha v\|_{0, \infty, \Omega} ; |\alpha| \leq m \}.$$

Using [4] for instance, one can easily prove

Lemma 7 Assume that Hypothesis H.2 holds. Then, there exists a constant  $c > 0$  independent of  $K \in \mathcal{T}_h$  such that

$$(5.2) \quad \|v - r_K v\|_{0, K} \leq c(h(K))^{k+1} \|v\|_{k+1, K} \quad \text{for all } v \in H^{k+1}(K),$$

$$(5.3) \quad \|v - r_K v\|_{0, K'} \leq c(h(K))^{k+3/2} \|v\|_{k+1, \infty, K} \quad \text{for all}$$

$v \in W^{k+1, \infty}(K)$  where  $K'$  is any side of  $\partial_+ K$



We are now able to prove

Theorem 6 Assume that all the elements  $K \in \mathcal{C}_h$  are rectangles,  
that  $\hat{P} = Q_k$  and that Hypothesis H.2 holds. Assume in addition,  
that the solution  $u$  of problem (1.3) belongs to  $H^{k+2}(\Omega)$ . Then,  
there exists a constant  $c > 0$  independent of  $h$  such that for all  
 $i = 1, \dots, I$

$$(5.4) \quad \|u_h - u\|_{0, \Omega_i} \leq Ch^{k+1} \|u\|_{k+2, \Omega_i},$$

$$(5.5) \quad \left( \int_{\partial_+ \Omega_i} (\mu n_x + \nu n_y) (u_h - u)^2 ds \right)^{1/2} \leq Ch^{k+1} (\|u\|_{k+2, \Omega_i} + \|u\|_{k+1, \infty, \Omega_i})$$

Proof For any  $K \in \mathcal{C}_h$ , we now define

$$(5.6) \quad \eta_h = \begin{cases} 0 & \text{on } \partial_- K \cap \Gamma_- , \\ \text{outward trace of } r_h u & \text{on } \partial_- K - (\partial_- K \cap \Gamma_-). \end{cases}$$

We start from equation (4.1) with  $v = r_h u$ ,  $\eta = \eta_h$ . The corresponding right hand side may be written in the form

$$(5.7) \quad X_K(u, u_h - r_h u) = Z_K(u, u_h - r_h u) + \int_K \sigma(u - r_h u) (u_h - r_h u) dx dy$$

where

$$(5.8) \quad \left\{ \begin{aligned} Z_K(u, w) &= \int_{\partial_+ K} (\mu n_x + \nu n_y) (u - r_h u) w ds + \int_{\partial_- K} (\mu n_x + \nu n_y) (u - \eta_h) w ds - \\ &- \int_K (u - r_h u) \left( \mu \frac{\partial w}{\partial x} + \nu \frac{\partial w}{\partial y} \right) dx dy . \end{aligned} \right.$$

We now use the following essential lemma which will be proved later.

Lemma 8 With the same assumptions as in Theorem 6, there exists a constant  $c > 0$  independent of  $K \in \mathcal{T}_h$  such that for all  $w \in Q_k$

$$(5.9) \quad |z_K(u, w)| \leq C(h(K))^{k+1} \|u\|_{k+2, K} \|w\|_{0, K}$$

Using (5.2), (5.7) and (5.9), we obtain for all  $K \in \mathcal{T}_h$

$$(5.10) \quad |X_K(u, u_h - r_h u)| \leq c_1 h^{k+1} \|u\|_{k+2, K} \|u_h - r_h u\|_{0, K}$$

Thus, combining (4.1) with  $v = r_h u$ ,  $\eta = \eta_h$ , (5.10) and using (3.7), we get

$$\left\{ \begin{array}{l} \frac{1}{2} \int_{\partial_+ K} (\mu n_x + \nu n_y) (u_h - r_h u)^2 ds + \alpha \|u_h - r_h u\|_{0, K}^2 \leq \\ \leq -\frac{1}{2} \int_{\partial_- K} (\mu n_x + \nu n_y) (\xi_h - \eta_h)^2 ds + c_1 h^{k+1} \|u\|_{k+2, K} \|u_h - r_h u\|_{0, K} \end{array} \right.$$

Summing over all the elements  $K_j$ ,  $1 \leq j \leq i$ , we obtain

$$(5.11) \quad \begin{aligned} & \frac{1}{2} \int_{\partial_+ \Omega_i} (\mu n_x + \nu n_y) (u_h - r_h u)^2 ds + \alpha \|u_h - r_h u\|_{0, \Omega_i}^2 \leq \\ & \leq c_1 h^{k+1} \|u\|_{k+2, \Omega_i} \|u_h - r_h u\|_{0, \Omega_i} \end{aligned}$$

Thus, the estimate (5.4) and (5.5) are simple consequences of inequality (5.11) and Lemma 7.

Proof of Lemma 8 Consider a rectangle  $K \in \mathcal{T}_h$  with vertices A, B, C, D (cf. Fig. 5). Let us denote by  $\Delta x$  (resp.  $\Delta y$ ) the length of the side AB (resp. BC). We may write

$$(5.12) \quad z_K(u, w) = \mu z_{K,x}(u, w) + \nu z_{K,y}(u, w)$$

with

$$z_{K,x}(u, w) = \int_D^A (u - r_h u) w \, dy - \int_C^B (u - \eta_h) w \, dy - \int_K (u - r_h u) \frac{\partial w}{\partial x} \, dx dy ,$$

$$z_{K,y}(u, w) = \int_B^A (u - r_h u) w \, dx - \int_C^D (u - \eta_h) w \, dx - \int_K (u - r_h u) \frac{\partial w}{\partial y} \, dx dy .$$

By using the one-to-one correspondence  $v \rightarrow \hat{v} = v \circ F_K$ , we get :

$$(5.13) \quad z_{K,x}(u, w) = \frac{\Delta y}{2} \hat{z}_x(\hat{u}, \hat{w})$$

with

$$\begin{aligned} \hat{z}_x(\hat{u}, \hat{w}) &= \int_{-1}^{+1} (\hat{u}(1, \hat{y}) - \hat{r}\hat{u}(1, \hat{y})) \hat{w}(1, \hat{y}) \, d\hat{y} - \\ &\quad - \int_{-1}^{+1} (\hat{u}(1, \hat{y}) - \hat{\eta}(\hat{y})) \hat{w}(1, \hat{y}) \, d\hat{y} - \int_{-1}^{+1} \int_{-1}^{+1} (\hat{u} - \hat{r}\hat{u}) \frac{\partial \hat{w}}{\partial \hat{y}} \, d\hat{x} d\hat{y} , \end{aligned}$$

where  $\hat{r}\hat{u}$  is the polynomial of  $Q_k$  which interpolates  $\hat{u}$  at the points  $\hat{a}_{ij}$ ,  $1 \leq i, j \leq k+1$ , and where  $\hat{\eta}$  is the polynomial of degree  $\leq k$  which interpolates the function  $\hat{y} \rightarrow \hat{u}(-1, \hat{y})$  at the points  $\theta_i$ ,  $1 \leq i \leq k+1$ .

Clearly

$$\hat{z}_x(\hat{u}, \hat{w}) = 0 \quad \text{for all } \hat{u}, \hat{w} \in Q_k .$$

Now, when  $\hat{u} = \hat{x}^{k+1}$ , we have

$$\hat{u}(1, \hat{y}) = \hat{r}\hat{u}(1, \hat{y}) = 1, \quad \hat{u}(-1, \hat{y}) = \hat{\eta}(\hat{y}) = (-1)^{k+1} .$$

Moreover,  $\hat{r}\hat{u}$  does not depend on  $\hat{y}$  and then, for all  $\hat{w} \in Q_k$ , the function  $\hat{x} \rightarrow (\hat{u} - \hat{r}\hat{u})(\hat{x}) \frac{\partial \hat{w}}{\partial \hat{x}}(\hat{x}, \hat{y})$  is a polynomial of degree  $\leq 2k$

which vanishes at the  $(k+1)$  Gauss-Radau points  $\theta_i$ . Therefore,

$$\int_{-1}^{+1} (\hat{u} - \hat{r}\hat{u}) \hat{w} \, d\hat{x} = 0 \quad \text{for all } \hat{w} \in Q_k .$$

Thus, when  $\hat{u} = \hat{x}^{k+1}$ , we get

$$\hat{z}_x(\hat{u}, \hat{w}) = 0 \quad \text{for all } \hat{w} \in Q_k .$$

On the other hand, when  $\hat{u} = y^{k+1}$ ,  $\hat{r}\hat{u}$  is independent of  $\hat{x}$  so that we obtain by integration by parts

$$\begin{aligned} \int_{-1}^{+1} \int_{-1}^{+1} (\hat{u} - \hat{r}\hat{u}) \frac{\partial \hat{w}}{\partial \hat{x}} \, d\hat{x} d\hat{y} &= \int_{-1}^{+1} (\hat{u}(1, \hat{y}) - \hat{r}\hat{u}(1, \hat{y})) \hat{w}(1, \hat{y}) \, d\hat{y} - \\ &- \int_{-1}^{+1} (\hat{u}(1, \hat{y}) - \hat{r}\hat{u}(1, \hat{y})) \hat{w}(1, \hat{y}) \, d\hat{y} . \end{aligned}$$

This give again

$$\hat{z}_x(\hat{u}, \hat{w}) = 0 \quad \text{for all } \hat{w} \in Q_k .$$

Therefore, we have proved that

$$\hat{z}_x(\hat{u}, \hat{w}) = 0 \quad \text{for all } \hat{u} \in P_{k+1} \quad \text{and all } \hat{w} \in Q_k .$$

Then, for fixed  $\hat{w} \in Q_k$ , the linear functional  $\hat{u} \rightarrow \hat{z}_x(\hat{u}, \hat{w})$  is continuous over  $H^{k+2}(\hat{K})$  with norm  $\leq c_1 \|\hat{w}\|_{0, \hat{K}}$  and vanishes over  $P_{k+1}$ . By the Bramble-Hilbert lemma in the form given in [4, Lemma 6], we get for all  $\hat{u} \in H^{k+2}(\hat{K})$  and all  $\hat{w} \in Q_k$

$$|\hat{z}_x(\hat{u}, \hat{w})| \leq c_2 |\hat{u}|_{k+2, \hat{K}} \|\hat{w}\|_{0, \hat{K}} .$$

Going back to the element  $K$  by using the correspondence

$\hat{v} \rightarrow v = \hat{v} \circ F_K^{-1}$  and (5.13), we obtain for all  $u \in H^{k+2}(K)$  and all  $w \in Q_k$

$$(5.14) \quad |z_{K,x}(u,w)| \leq c_3 (h(K))^{k+1} \|u\|_{k+2,K} \|w\|_{0,K}$$

Likewise, we get

$$(5.15) \quad |z_{K,y}(u,w)| \leq c_4 (h(K))^{k+1} \|u\|_{k+2,K} \|w\|_{0,K}$$

Then, combining (5.12), (5.14) and (5.15), we obtain the desired inequality (5.9).  $\square$

Note that the error estimates of Theorem 6 are now optimal in the exponent of the parameter  $h$ . However, as the one-dimensional results of § 1 suggest, we conjecture that, for any rectangle  $K \in \mathcal{T}_h$ , there exist some points of  $\partial_+ K$  where even more precise error bounds hold. Unfortunately, we have not been able to prove the existence of such points.

REFERENCES

- [ 1 ] Axelsson, O., " A class of A-stable methods " , B.I.T. 9 (1969), 185-199.
- [ 2 ] de Boor, C., and B. Swartz " Collocation at Gaussian points " , SIAM J. Numer. Anal. 10 (1973) 582-606.
- [ 3 ] Butcher, J.C., " Implicit Runge-Kutta processes " , Math. Comp. 18 (1964) , 50-64.
- [ 4 ] Ciarlet, P.G., and P.A. Raviart " General Lagrange and Hermite interpolation in  $\mathbb{R}^n$  with applications to finite element methods " , Arch. Rat. Mech. Anal. 46 (1972), 177-199.
- [ 5 ] Ciarlet, P.G., and P.A. Raviart " Interpolation theory over curved elements, with applications to finite element methods " , Comp. Meth. Appl. Mech. Eng. 1 (1972), 217-249.
- [ 6 ] Ciarlet, P.G., and P.A. Raviart " The combined effect of curved boundaries and numerical integration in isoparametric finite element methods " , The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations (A.K. Aziz, Editor), 409-474, Academic Press, New-York, 1972.
- [ 7 ] Cruzeix, M., Thesis , to appear.
- [ 8 ] Friedrichs, K.O., " Symmetric positive differential equations " , Comm. Pure Appl. Math. 11 (1958), 333-418.
- [ 9 ] Hulme, B.L., " Discrete Galerkin and related one-step methods for ordinary differential equations " , Math. Comp. 26 (1972), 881-891.
- [ 10 ] Kaper, H.G., G.K. Leaf and A.J. Lindeman " Application of finite element techniques for the numerical solution of the neutron transport and diffusion equations " Proc. Conf. on Transport Theory , 2nd Conf. 710107 , Los Alamos (1971).
- [ 11 ] Lesaint, P., " Finite element methods for symmetric hyperbolic equations " , Numer. Math. 21 (1973) , 244-255.
- [ 12 ] Lesaint, P., " Finite element methods for the transport equation " , to appear in RAIRO , série Mathématiques.
- [ 13 ] Lesaint, P., Thesis , to appear.
- [ 14 ] Lesaint, P., and J. Gérin-Roze " Isoparametric finite element methods for the neutron transport equation " , to appear.
- [ 15 ] Miller, W.F. Jr, E.E Lewis, and E.C. Rossow " The application of phase-space finite elements to the two-dimensional transport equation in x-y geometry " , to appear in Nucl. Sci. Eng.
- [ 16 ] Ohnishi, T., " Application of finite element solution technique to neutron diffusion and transport equations " , Proc. conf. on New Developments in Reactor Mathematics and Applications , CONF - 710302, Idaho Falls (1971).
- [ 17 ] Reed, W.H., and T.R. Hill " Triangular mesh methods for the neutron transport equation " , to appear in Proc. Amer. Nucl. Soc.



Université Paris VI  
et

Centre National de la Recherche Scientifique

LABORATOIRE ASSOCIÉ 189  
**ANALYSE NUMÉRIQUE**

Tour 55. 65. 5<sup>ème</sup> Étage  
4, place Jussieu  
75230 - PARIS CEDEX 05  
Tél. : 336-25-25, poste 42-98

TRAVAUX DU LABORATOIRE ASSOCIÉ 189

Année Universitaire 1973-1974.

- M. CROUZEIX, P.A. RAVIART. Conforming and Nonconforming Finite Element Methods for Solving the Stationary Stokes Equations . I  
(Ref. 73001).
- M. CROUZEIX, J.M. THOMAS. Éléments finis et problèmes elliptiques dégénérés.  
(Ref. 73002).
- H. BREZIS, G. DUVAUT . Ecoulements avec sillages autour d'un profil symétrique sans incidence.  
(Ref. 73003).
- V. GIRAULT. Theory of a Finite Difference Method on irregular Networks for Solving General Elliptic Dirichlet Boundary Value Problems of Order Two.  
(Ref. 73004) .
- P.A. RAVIART. Méthode des éléments finis. Cours de D.E.A analyse numérique rédigé par J.M. THOMAS.  
(Ref. 73005) .
- P.A. RAVIART. The Use of Numerical Integration in Finite Element Methods for Solving Parabolic Equations.  
(Ref. 73006) .
- A. PERRONNET. Décompositions du problème de Stokes, stationnaire et évolutif par dualisation de certaines contraintes. Rédigé en collaboration avec M. BONNET.  
(Ref. 73007) .



J.P. BOURGUIGNON, H. BREZIS. Remarks on the Euler Equation.

(Ref. 73008) .

J.L. LIONS, P.A. RAVIART. Séminaires d'analyse numérique.(1972-1973)

(Ref. 73009) .

J.L. LIONS, H. BREZIS. Séminaires sur les équations aux dérivées partielles.

(1972-1973) (Ref. 73010) .

H. BREZIS. Solutions with Compact Support of Variational Inequalities.

(Ref. 73010) .

P.A. RAVIART. Finite Element Methods for Solving the Stationary Stokes and Navier-Stokes Equations.

(Ref. 73011) .

P. FAURRE. Identification et filtrage statistique. D.E.A. analyse numérique

(1973-1974) (Ref. 73012) .

P.G. CIARLET. Sur l'élément de Clough et Tocher.

(Ref. 73013) .

P.G. CIARLET. Quelques méthodes d'éléments finis pour le problème d'une plaque encastrée.

(Ref. 73014) .

LABORATOIRE ASSOCIÉ 189  
ANALYSE NUMÉRIQUE

Tour 55. 65. 5<sup>ème</sup> Étage  
4, place Jussieu  
75230 - PARIS CEDEX 05  
Tél. : 336-25-25, poste 42-98

TRAVAUX DU LABORATOIRE ASSOCIÉ 189

Année Universitaire 1974-1975.

C. GERHARDT. Existence and Regularity of Capillary Surfaces.

(Ref. 74001).

P.G. CIARLET, P.A. RAVIART. A Mixed Finite Element Method for the Biharmonic Equation.

(Ref. 74002).

F. MURAT, SIMON. Quelques résultats sur le contrôle par un domaine géométrique.

(Ref. 74003).

J.L. LIONS, P.A. RAVIART. Séminaires d'analyse numérique. (1973-1974).

(Ref. 74004).

L. TARTAR. Interpolation. Cours de III cycle rédigé par F. MURAT.

(Ref. 74005).

P.G. CIARLET, P.A. RAVIART. La méthode des éléments finis pour les problèmes aux limites elliptiques. Chap. I.

(Ref. 74006).

M.F. BIDAUT. Théorèmes d'existence et d'existence "en général" pour des problèmes de contrôle optimal.

(Ref. 74007).